ISSN 2348 – 9928
IJAICT Volume 3, Issue 4, April 2016 Doi:01.0401/ijaict.2014.05.11 Published on 05 (4) 2016

A SURVEY ON THE PROPAGATION OF WEB VIDEOS

Mr. K. Vishnu
PG Scholar,
Computer Science and Engineering,
Anna University Regional Centre,
Coimbatore, Tamilnadu, India

Mr. V. Yuvaraj
Teaching Assistant,
Computer Science and Engineering,
Anna University Regional Centre,
Coimbatore, Tamilnadu, India

Mr. J. Dhandapani
PG Scholar,
Computer Science and Engineering,
Anna University Regional Centre,
Coimbatore, Tamilnadu, India

Abstract— ***Our aim is to take a survey about the propagation of web videos that is by estimating two things. One is the estimation of video influences (i,e) how the video plays a role in everyone's life, it may be a public issue or a private issue. Second is the estimation of origin of videos (i,e) how the video get starting origination into outside world. It may be get originated from a popular site or from some particular person's website and get popularity. As a result, it is becoming more difficult to determine how the propagation took place - was the video a piece of original work that was intentionally uploaded to its major hosting site by the authors, or did the video originate from some small site then reached the sharing site after already getting a good level of popularity, or did it originate from other places in the cyberspace but the sharing site made it popular. It is not an easy thing; nowadays this issue plays an important role. Thus we take these both issues into account and collect different methods which and all gives a better solution to this.***

Keywords— Origination, Influence, Online Videos, Social Network, Cyber Space.

I. INTRODUCTION

Web mining is the one of the application of data mining techniques to discover patterns from the web. They can be divided into three different types, which are Web usage mining, Web content mining and Web structure mining. Web usage mining is the process of extracting useful information from server logs e.g. use Web usage mining is the process of finding out what users are looking for on the Internet. Web structure mining is the process of using graph theory to analyze the node and connection structure of a web site. Web content mining is the mining, extraction and integration of useful data, information and knowledge from Web page content [1].

Video is a visual presentation, typically a moving picture which is accompanied by sound. Along with the technical progress in Internet technology and broadband connections a booming number of webmasters include videos in their websites [2]. As Internet users are less and less willing to read text, video is an excellent way to bring one's message across. Be it a testimonial or demonstration of a software, a filmed lecture or just something funny or intriguing, people enjoy watching movies. Web Videos are either filmed videos taken with a camera (click here to view an example) or are screen-capture video where the action on a computer screen is recorded (click here for an example) [3].

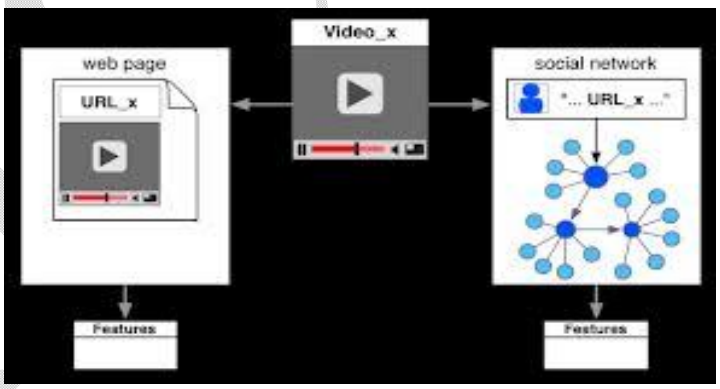

Fig 1: Propagation of online videos

As shown in the above figure the video is get from the web page and the features are extracted and then it is propagated in the social network. The URL of the video is detected and the origin and influence of the video is estimated [4]. Online social networks have become widespread and increasingly popular.

A common activity on these networks is sharing of content. Since the advent of social media hosting sites, users have vast repositories of external content to share. For content creators interested in the impact of their work, prediction of the propagation of their content is a valuable asset. In this thesis we work on predicting the spread of online video content on the social network Twitter. We take two perspectives in extracting information about videos and how they are shared.

A feature-based model identifies the content and Meta data of a video. Our propagation model describes how the sharing of a

video develops over time. By indicating correlations between features of the video content on one side and characteristics of propagation patterns on the other, we show how videos spread differently depending on their category, length and age. We use these findings to form a non-linear prediction model that forecasts social propagation of a video based on its features and an initial propagation period. This statistic shows the influence of online videos on the way US consumers thought about, planned, or booked travel as of June 2013. During the survey, 64 percent of business travellers Stated that online videos influenced where they decided to travel and 58 percent of leisure travellers said the same. In the below section we discuss about some of the techniques which are used for the estimation techniques of online videos and continued by its conclusion.

II. SURVEY OF EXISTING METHODS

2.1 Label Propagation through Linear Neighbourhoods

A novel semi-supervised learning approach is proposed here which was based on a linear neighbourhood model. The data point was linearly reconstructed from its neighbourhood. Also, the Linear Neighbourhood Propagation (LNP) algorithm was proposed to transmit the labelled points to the dataset using the linear neighbourhoods with sufficient softness [5] .LNP reduces the graph by a series of coincided linear neighbourhood patches. The standard quadratic programming procedure was used to solve the edge weight in each patch. These edge weights are combined together to form a weight matrix called Laplacian matrix. An easiest method was derived to extend the LNP to out of model data. Finally the experiments on digits, synthetic data and text classification are evaluated to show the electiveness of LNP.

2.2 Recommending Social Events from Mobile Phone Location Data

A city suggests a variety of social events a day, and it is difficult for dwellers to make choices. The mobile phones and recommender systems together can change the way one deals with such abundance. Mobile phones with locating technology are now available which makes easy for people to broadcast their location. The recommender system was used to identify the recommend events based on people's movements. The cold-start users, who have no location history and cannot, receive any recommendations. To identify the recommendations to cold-start users the large number of mobile phone users is collected from two datasets. A various algorithms are performed on these data for recommending social events. The experimental results are finally evaluated to show the effectiveness of these algorithms.

2.3 Authoritative Sources in a Hyperlinked Environment

The hyperlinked environment was a source of information which consists of the content of environment. A various algorithmic tools are developed to extract the information from the link structures of such environments. The main problem to identify the within the framework is the distillation of wide search topics, through the discovery of "authoritative" information sources on such topics [6].

An algorithmic formulation was tested of the conception of authority which was based on the relationship between a group of relevant authoritative pages and hub pages that connect them together in the link structure. These formulations had connected to the eigenvectors of several matrices associated with the link graph; these connections are used to motivate the extra heuristics for link-based analysis [7].

2.4 Support-Vector Networks

The grouping problem can be solved by the support-vector network which was a new learning machine. It implements the idea of each input vectors are nonlinearly mapped to a very high- dimension feature space where a linear decision surface was constructed. The properties of the decision surface provide high generalization ability of the learning machine [8]. The previous concept of the support-vector network was implemented with the separable training data without any errors. But here this concept was applied with non-separable training data. The performance of the support-vector network was compared with several learning algorithms which are all the part of Optical Character Recognition.

2.5 Person Re-Identification by Descriptive and Discriminative Classification

Person re-identification means identifying a single person across spatially disjoint cameras was an important task in visual surveillance. The previous methods either try to found an appropriate description of the appearance or absorb a discriminative model. Since these methods had some disadvantages so to overcome these drawbacks the proposed approach was demonstrated on two datasets, where we show that the combination of a generic descriptive statistical model and a discriminatively learned feature-based model reaches better results than the individual models. Also, the state-of-the-art on a

publicly contains benchmark dataset was compared with the proposed approach [9].

2.6 The Joint Inference of Topic Diffusion and Evolution in Social Communities

A Novel probabilistic model was proposed here to solve the joint inference problem which consists of the textual documents, social influences, and topic progress in a unified way. The evolution of an arbitrary topic and expose the latent diffusion paths of the corresponding topic in a social community was tracked first. Then based on the diffusion and the evolution of the topic, a mixture model was introduced to model the generation of text. The entire diffusion process was regularized through a Gaussian Markov Random Field. The experimental results on joint interference based on the diffusion and evolution of the topic and the proposed probabilistic model shows the better performance than the previous methods.

2.7 I Tube, You Tube, Everybody Tubes: Analysing the World's Largest User Generated Content Video System

The analysis of you tube which was the world's largest User Generated Content (UGC) Video system was presented here to overcome the drawback of UGC systems. The large data's are collected from the you tube to provide the complete study of YouTube and other similar UGC systems. The understandings of the potential for very efficiently UGC video systems are also provided, for example utilizing P2P techniques. Finally, the opportunities to leverage the latent demand for function videos that are not reached today due to some system scarcity distortions are discussed. The experimental results are essential to understand the UGC systems and it provides useful information to ISPs, site administrators, and content owners with major commercial and technical implications.

2.8 Extracting Shared Subspace for Multi-label Classification

A general framework for extracting shared structures in multi-label classification was proposed here. A common subspace is considered to be shared among multiple labels in this framework. The Eigen value problem was solved to get the optimal solution to the proposed formulation. An efficiently algorithm was developed to solve the high-dimensional problems. The experiments are conducted on eleven multi topic web page categorization tasks and it results proves the effectiveness of the proposed formulation in comparison with several representative algorithms.

2.9 Open Domain Event Extraction from Twitter

This paper defines TwiCal which means the first open-domain event extraction and classification system for Twitter. The extraction of an open domain calendar of significant events from Twitter was exactly possible which was determined here. Also the novel method for discovering important event categories and classifying extracted events based on latent variable models are presented here. This approach achieves a 14% increase in maximum F1 over a supervised baseline by leveraging large volumes of unlabelled data.

III. CONCLUSION

Online videos are so popular nowadays that they begin to change people's way of daily entertainment greatly. The study of how online videos propagate and how influential they are outside a video sharing site is an increasingly significant research problem. The identification of an online video's origin and propagation patterns, from the video sharing site's perspective, is crucial to its business models as well to its partner's decision making for their marketing strategies. Thus in this paper we present a study of various techniques which and all make worth for our estimations. From this survey we determine if a popular online video originated from the video sharing site, or from somewhere else of the Internet and also we determine how it got popular through its analyzing life cycle.

References

[1] Corinnacortes, and Vladimirvapnik, "Support-Vector Networks", 1995.

[2] Jon M. Kleinberg, "Authoritative Sources in a Hyperlinked Environment",1998.

[3] Fei Wang, and feiw Changshui Zhang, "Label Propagation Through Linear Neighborhoods", International Conference on Machine Learning, 2006.

[4] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon, "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System", 2007.

[5] Shuiwang Ji, Lei Tang, Shipeng Yu, and Jieping Ye, "Extracting Shared Subspace for Multi-label Classification", 2008.

[6] Daniele Querciaxy, Neal Lathiaz, Francesco Calabresey, Giusy Di Lorenzoy, and Jon Crowcroft, "Recommending Social Events from Mobile Phone Location Data", IEEE, 2010.

[7] Cindy Xide Lin, Qiaozhu Mei, Jiawei Han, Yunliang Jiang, and Marina Danilevsky, "The Joint Inference of Topic Diffusion and Evolution in Social Communities", IEEE, 2011.

[8] Martin Hirzer, Csaba Beleznai, Peter M. Roth, and Horst Bischof, "Person Re-Identification by Descriptive and Discriminative Classification", 2011.

[9] Alan Ritter, Mausam, Oren Etzioni, and Sam Clark, "Open Domain Event Extraction from Twitter", 2012.